Confirming the structure of negative beliefs about psychosis and bipolar disorder: A confirmatory factor analysis study of the Personal Beliefs about Experience Questionnaire (PBEQ) and Personal Beliefs about Illness Questionnaire (PBIQ).

Abstract

Objectives: Negative beliefs about psychosis and other mental health difficulties may contribute to depression and distress in individuals with these experiences. The Personal Beliefs about Experience Questionnaire (PBEQ) and Personal Beliefs about Illness Questionnaire (PBIllQ) are two widely used measures of these beliefs. It is currently uncertain how the items on these measures map onto different underlying factors. The current study therefore aimed to test the factor structure of these two measures.

Methods: Confirmatory Factor Analysis (CFA) was used to test three alternative, pre-specified, factor structures for the PBIllQ and PBEQ in a sample of individuals diagnosed with bipolar disorder (n = 202) and a sample of individuals with experiences of psychosis (n = 362). Associations with depressive symptoms were also examined.

Result: A three-factor structure was supported for both measures, which included Negative Expectations/Appraisals (NEA), Internal Shame/Defectiveness (ISD) and External Shame (ES) factors. The NEA and ISD subscales also had consistent independent associations with depressive symptoms.

Conclusions: The results suggest that the PBIllQ and PBEQ may capture three distinct sets of negative beliefs in individuals with psychosis or bipolar disorder, and that these beliefs may have important consequences for subsequent difficulties in these populations such as depression. Both measures may be helpful in supporting assessment and formulation in clinical practice and in evaluating belief change in intervention trials. However, when used in these settings the three subscales identified in this study may be the most valid way of calculating scores on these measures.

Keywords: Psychosis; Bipolar Disorder; Factor Analysis; Psychometrics; Negative Beliefs

Practitioner Points

- Negative personal beliefs about the causes, meaning and consequences of psychosis and bipolar disorder are associated with greater distress and depression. Two related measures, the PBIllQ and PBEQ, have been developed to assess these beliefs.
- Our analyses suggest that scores on these questionnaires are best broken down into three subscales which capture perceptions of internal shame or defectiveness, general negative appraisals, and perceptions of external shame.
- These subscales may capture key underlying sets of negative beliefs within individuals with psychosis or bipolar disorder, which in turn impact upon well-being, such as being associated with greater depression.
- These subscales can be used in order to aid assessment and formulation within clinical practice but may also provide a valuable means of assessing changes in negative beliefs following interventions.

Confirming the structure of negative beliefs about psychosis and bipolar disorder: A confirmatory factor analysis study of the Personal Beliefs about Experience Questionnaire (PBEQ) and Personal Beliefs about Illness Questionnaire (PBIllQ)

Individuals with experiences of psychosis or bipolar disorder develop beliefs about the causes, consequences and meanings of these difficulties (Birchwood, Mason, MacMillan, & Healy, 1993). For example, a variety of beliefs regarding the cause of their difficulties have emerged from qualitative interviews with individuals diagnosed with schizophrenia and bipolar disorder (Clatworthy, Bowskill, Rank, Parham, & Horne, 2007; Sayre, 2000). Likewise those deemed at risk of developing psychosis report concerns around the meaning of their experiences and "going mad" (Byrne & Morrison, 2010). Internalized stigma represents another potential set of beliefs associated with mental illness, including perceptions of being different or being "unworthy" of social contact (Green, Hayes, Dickinson, Whittaker, & Gilheany, 2003). The current paper reports on a psychometric evaluation of two related measures designed to assess negative beliefs about psychosis and bipolar disorder.

Negative beliefs about psychotic experiences are clinically important since such beliefs are associated with recovery style, quality of life, depression, anxiety and hopelessness (Acosta, Aguilar, Cejas, & Gracia, 2013; Birchwood et al., 1993; Karatzias, Gumley, Power, & O'Grady, 2007; Pyle et al., in press; Stainsby, Sapochnik, Bledin, & Mason, 2010; Stowkowy, Perkins, Woods, Nyman, & Addington, 2014; Stowkowy, Perkins, Woods, Nyman, & Addington, in press; White, McCleery, Gumley, & Mulholland, 2007). Such beliefs may therefore represent one mechanism through which affective disturbances in individuals with psychotic experiences develop. Whilst there has been less research into these belief sets in bipolar disorder there is evidence of such beliefs being associated with treatment decisions and depression (Elnazer, Brannac, & Kingdon, 2014; Scott & Pope, 2002). An

understanding of the underlying structure of these beliefs is important for accurate conceptualisation and measurement.

The first measure, the Personal Beliefs about Illness Questionnaire (PBIllQ; Birchwood et al., 1993), was developed to assess negative beliefs in those diagnosed with psychotic disorders and bipolar disorder. The second measure, the Personal Beliefs about Experiences Questionnaire (PBEQ; Pyle et al., in press) was adapted from the PBIllQ with the aim of producing a measure of individuals' beliefs about psychotic experiences that would be applicable to those deemed "at-risk" of psychosis as well as those with a diagnosis of a psychotic disorder (Pyle et al., in press). Adaptions to the PBIllQ therefore included the exclusion of some items (e.g., those relating to psychiatric services and relapse), which may be less relevant to an at-risk population. The language was also altered to avoid potential stigma, so that the term "illness" was replaced with "experiences".

The current study aimed to test the factor structure of the PBIllQ and PBEQ by comparing three alternative models. Identifying the most suitable factor model is important not only in guiding how the measures are used in clinical and research settings (e.g., how subscales are created) but also in exploring the phenomenological architecture of individuals' beliefs about their psychotic experiences. As both measures have been used as outcomes in clinical trials (Morrison et al., 2012), the construct validity (which is partly assessed through factor analysis; O'Leary-Kelly & Vokurka, 1998) of these measures is important if meaningful inferences are to be made.

Based on their common origins and content a similar factor structure would be expected across both measures. Three alternative models are considered. The first model is the two-factor structure initially identified through Principal Components Analysis (PCA) by Pyle and colleagues (in press) of the PBEQ items. The first factor, Negative Appraisals of Experiences (NAE), reflects general negative perceptions and expectations regarding the

ability to cope with or get on with life in the face of psychotic experiences (e.g., "I am capable of very little as a result of my experiences"). The second factor, Social Acceptance of Experiences (SAE), reflects stigma and feelings of being abnormal or different (e.g., "I am embarrassed to talk about my experiences"). Whilst PCA supported this two-factor structure, it is notable that the SAE factor had only moderate internal consistency ($\alpha = .52$). The low reliability of the SAE subscale is problematic in that it may attenuate the strength of relationships and suggests the factor may not be very robust.

An alternative factor structure is provided by Birchwood and colleagues' original PBIllQ (1993), from which the PBEQ was developed. Within this scale, five factors were originally specified; perceptions of control over the illness ("My illness frightens me"), self as illness ("There is something about my personality that causes this illness"), negative expectations ("I will always need to be cared for by professional staff"), Stigma ("I am embarrassed by my illness") and social containment ("Society needs to keep people like me who have this illness, apart from everyone else"). This initial factor structure was derived from theoretical considerations but not confirmed via factor analysis.

A three factor model was also devised by the study authors based on the observation that a number of items relate to perceptions of internal shame (i.e., judging the self as unworthy or invalid; Gilbert, 2006) and defectiveness ("My experiences are a judgement on me"). We've labelled this the Internal Shame and Defectiveness (ISD) factor. Such a factor is consistent with evidence of defectiveness/shame schema occurring in individuals with psychosis (Addington & Tran, 2009; Fowler et al., 2006). Internal and external forms of shame may be distinct (Gilbert, 2006), and thus this ISD factor can be differentiated from items relating to External Shame (ES; "I am embarrassed to talk about my experiences", "I can talk to most people about my experiences [reversed]"). Thus the ISD factor may capture self-to-self relating and forms of internalised stigma whilst the ES factor may capture other-

to-self perceptions (Gilbert, 2006). The remaining items on both measures can be grouped into a general Negative Appraisals of Experiences (NAE) factor, where experiences are viewed as frightening, difficult to cope with and inhibiting opportunities.

Confirmatory Factor Analysis (CFA) was used to test and compare these three distinct factor models (2 vs. 5 vs. 3 factors) in a sample of individuals with experiences of psychosis or bipolar disorder. Previous research suggests that an individual's beliefs about their unusual or psychotic experiences will be a key predictor of affective disturbances in this population. Therefore, in addition to comparing the relative fit of these models, we also examined the cross-sectional relationships between the factors produced by these models and depressive symptoms. The PBEQ represents a slight adaption to the PBIllQ. In light of their similarity both measures would not be used for the same individuals. Hence we evaluated these measures in separate samples.

Study 1

The first study compares the three putative factor models for the PBIllQ in a sample of individuals diagnosed with bipolar disorder.

Method

Participants

Participants with a life time diagnosis of bipolar disorder were recruited from mental health services in Scotland to take part in a trial of Cognitive Interpersonal Therapy. DSM-IV (American Psychiatric Association, 2000) diagnoses of bipolar I disorder were confirmed by both the referring psychiatrist and by researchers via the Structured Clinical Interview for DSM-IV (First, Spitzer, Gibbon, & Williams, 1996). Exclusion criteria were non-English speakers, presence of organic brain disorder, significant intellectual disability or receipt of electro-convulsive therapy or psychotherapy beyond that delivered in the trial. The sample

consisted of n = 202 individuals (n = 98 male) aged between 17 and 65 years (M = 36.56, SD = 10.60). Ethnicity was not recorded. Ethical approval was obtained for the trial.

Measures

Personal Beliefs about Illness Questionnaire (PBIllQ). The PBIllQ is a 16-item measure of an individual's beliefs or appraisals of their mental health difficulties, and are rated on a four-point scale from "strongly disagree" to "strongly agree". Items relate to the perceived causes and consequences of having a diagnosis of a severe and enduring mental health problem.

Beck Depression Inventory (BDI-II; Beck, Steer, & Brown, 1996). This is a 21-item measure assessing depressive symptoms including sadness, pessimism and loss of pleasure over the previous two weeks using a four-point scale. The reliability, factor structure, concurrent and content validity of this measure has been supported (Wang & Gorenstein, 2013) and scores have been found to distinguish between different depressive and non-depressive episodes in bipolar disorder (Kumar, Rissmiller, Steer, & Beck, 2006). Internal reliability was α = .94 in the current sample.

Statistical Analysis

CFA were conducted in R R3.0.1 (R Core Team, 2014) using the LAVAAN package (Rosseel, 2012). As the PBEQ items were ordinal in nature, Mean and Variance Weighted Least Squares Estimation (WLSMV) was used based upon a polychoric correlation input matrix. This approach has been shown to perform well with ordinal data (Kline, 2011). Model fit was determined via the scaled goodness-of-fit statistic which tests the null-hypothesis of exact fit to the data. Model fit was also determined via the Confirmatory Fit Index (CFI; adequate fit > .90, good fit > .95), Tucker-Lewis Index (TLI; adequate fit > .90, good fit > .96) and the Root Mean Squared Residual (RMSEA; adequate fit < .08, good fit < .06 with the upper confidence interval < .08) based on recommendations by Kline (2011), Hu

& Bentler (1999) and Byrne (2001). Rates of missing data within this sample were very low (< 1% per variable) and removed on a casewise basis prior to analysis (final CFA n = 194). Two of the PBIllQ items are reverse scored, and these were therefore reflected so that all inter-correlations between items were positive.

Results

Model Comparisons

Each of the three alternative factor models were estimated in turn and the fit statistics are reported in Table 1. Model 1 is the two-factor model outlined by Pyle and colleagues (in press), Model 2 is the three factor model suggest within this paper, whilst Model 3 is the five-factor model based upon the original paper by Birchwood and colleagues (1993). Notably all models converged normally except for Model 3, which produced an inadmissible factor correlation between Social Containment and the other factors ($r > 1$). This likely reflects model misspecification, and in particular the small number of indicators for the SC factor and its close conceptual relationship with other factors. Hence results for Model 3 should be interpreted with caution. Model 2 (3 factor) was the best fitting of the three alternative models. However, the overall fit of this model still fell below our criteria for good fit.

TABLE 1 AROUND HERE

Model Development

Modification indices for Model 3 suggested allowing the residuals associated with items one ("I will always need to be cared for by professional medical staff") and five ("Because of my illness I have to rely on psychiatric services"), to correlate. This correlation makes conceptual sense, sharing a common theme of psychiatric service involvement. Allowing these residuals to correlate led to a better fitting model (Model 4; this model is nested within model 2, allowing a direct comparison), scaled $\Delta X^2 = 60.48$, $p < .01$. Allowing a further residual correlation for items 14 ("There must be something about my personality

that causes me to be what I am") and 16 ("There is something strange about me that causes my illness") led to a further improvement in the model (Model 5), scaled $\Delta X^2 = 19.06$, $p < .01$. Twenty-eight out of 60 standardized implied residuals for this model fell above .10 (multiple values above .10 may be indicative of misspecification; Kline, 2011). The parameters associated with this final model are reported in Table 2. All standardized loadings were above .40 meeting recommendations for minimum factor loadings size (Costello & Osborne, 2005).

TABLE 2 AROUND HERE

The data for model 5 were checked for outlying cases that may have been having an undue influence upon the model, using Cook's distances estimated via the FAOUTLIER package for R, which can take the ordinal nature of the data into account (Chalmers, 2013). The largest Cook's distance was 0.013, falling below the recommended cut-off of $4/N = 0.02$ for potential influential cases (Bollen & Jackman, 1990), suggesting that no individual cases were having an undue influence upon the model.

Subscale Reliability and Validity

Three subscale scores were created by summing the items that loaded onto the three factors in Model 5. Internal consistencies and descriptive statistics for the three subscales are reported in Table 3. Whilst the ISD and NEA had acceptable internal reliability, reliability was lower for the ES subscale. This no doubt reflects the small number of items loading onto this factor (n = 2). The three subscales demonstrated large to moderate correlations with depressive symptoms (Table 3). Within a multiple regression analysis, the ISD and NAE subscales, but not the ES subscale, made significant independent contributions to the prediction of depressive symptoms (see semi-partial correlations in Table 3), accounting for 45% of the variance, $F(3, 186) = 50.77$, $p < .01$. Residuals were homoscedastic and normally

distributed. Equivalent findings emerged after three cases with large standardized residuals (> 2.5) were dropped.

TABLE 3 AROUND HERE

Discussion

The three factor model was the best fitting for the PBIllQ. The overall fit of the final model still fell below our criteria for good fit, and there were a number of larger standardized fitted residuals, suggesting a degree of possible misspecification. Allowing for further correlated residuals may have improved fit but would run the risk of over-fitting. There was no obvious indication of an alternative, better fitting, factor structure from the results. Consequently, the three factor structure proposed here may be preferable for future use of the PBIllQ, but further psychometric evaluation of this measure may be warranted. The ES subscale has limited internal reliability due to the small number of items and did not make any independent contribution to the prediction of depression. Future users of the PBIllQ may therefore wish to focus on the ISD and NAE subscales.

Notably, a revised version of the PBIllQ has since been developed (Birchwood, Jackson, Brunet, Holden, & Barton, 2012) that has received a more robust psychometric development. This revised form of the PBIllQ excludes a number of earlier items and includes a number of new items, so that it differs in content from the PBEQ scale. However, the theorized distinction between factors has not been confirmed in this new measure (in this study factor analysis was only applied to subsets of the items so the emergence of the hypothesised factors from the data was not tested). In the absence of empirical support for the hypothesised distinctions between subscales scores based on this revised measure may lack validity. Future research could explore whether the three factors supported by this study also emerge in the revised PBIllQ.

Study 2

In this second study, the three competing factor models were tested for the PBEQ in a sample of individuals with experiences of psychosis. As the PBEQ excludes some items from the original PBIllQ it was impossible to test the social containment factor (part of the five-factor structure proposed by Birchwood et al., 1993), as only one item from this putative factor remains in the PBEQ. Therefore, this item ("My experiences may mean that I should be kept away from others") was included in the stigma factor, leading to a four factor structure.

Method

Participants

Data were collected as part of 1) the Assessment of Cognitive Therapy Instead of Neuroleptic (ACTION) Trial and 2) the Early Detection and Intervention Evaluation (EDIE) 2 Trial. Participants were eligible for the ACTION trial if they were aged between 16-65, either met criteria for entry into Early Intervention in Psychosis Service or had an ICD-10 diagnosis on the schizophrenia spectrum and had either been offered anti-psychotic medication and refused or had discontinued medication for 6 months or more. On entry into the trial all participants were experiencing psychotic symptoms as measured by the Positive and Negative Syndrome Scale (Kay, Fiszbein, & Opler, 1987), which was defined as a score of 4 or more on delusions or hallucinations or a score of 5 or more on suspiciousness, grandiosity or conceptual disorganisation. Full details regarding the recruitment procedure and referral sources for the ACTION trial can be found in Morrison and colleagues (2013). Participants were eligible for EDIE 2 if they were aged between 14 and 35 years old and met criteria for an At Risk Mental State (ARMS) on the Comprehensive Assessment of At Risk Mental States (CAARMS; Yung et al., 2005). Full details regarding the recruitment procedure and referral sources can be found in Morrison and colleagues (2011). Ethical approval was obtained for both trials and all participants provided full informed consent

before entry into the trial. All measures were completed at baseline assessment. The overall sample consisted of n = 362 individuals (n = 143 female), aged between 14 and 64 years (*M* = 22.96, SD = 7.99). They had on average 12.96 years of education (*SD* = 2.31). Participants were predominantly White (87.02%, *n* = 315), with 3.87% being Black (*n* = 14), 2.76% South Asian (*n* = 10) and 3.87% other ethnicity (*n* = 14), with some values missing (*n* = 9).

Measures

Personal Beliefs about Experiences Questionnaire (PBEQ). The PBEQ is a 13-item measure of an individual's beliefs or appraisals of their psychotic experiences, adapted from the PBIllQ for use with those at risk of developing psychosis as well as those experiencing frank psychosis. Items relate to the perceived causes and consequences of psychosis. As with the PBIllQ, items are rated on a four point scale from "strongly disagree" to "strongly agree". A copy is available in Appendix II.

Beck Depression Inventory for Primary Care (BDI-PC; Beck, Guth, Steer, & Ball, 1997). This is a seven-item measure assessing depressive symptoms including sadness, pessimism and loss of pleasure over the previous two weeks using a four-point scale. This measure has good sensitivity (97%) specificity (99%) in predicting major depression in a primary care sample (Steer, Cavalieri, Leonard, & Beck, 1999) and has been used previously in individuals with experiences of psychosis (Taylor et al., in press). Internal reliability was α = .81 in the current sample.

Statistical Analysis

This was as for Study 1. Missing data were minimal but greater than for Study1, ranging from 14.09% to 17.13% across the PBEQ items. The most prominent pattern of missing data was total non-response of the items (50 cases). Little's Missing Completely at Random test was also non-significant, X^2 (144) = 110.57, p = .98, suggesting data were missing completely at random. Missing data could therefore be excluded from the analysis

casewise without bias (Schafer & Graham, 2002), with n = 288 complete cases remaining. Two of the PBEQ items are reverse scored, and these were therefore reflected so that all inter-correlations between items were positive.

Results

Model Comparisons

The model fit statistics are reported in Table 4. Model 2 (3 factor) was the best fitting of these three alternative models. However, the overall fit of this model still fell below our criteria for good fit.

TABLE 4 AROUND HERE

Model Development

Modification indices for Model 3 suggested that allowing the residuals associated with items seven ("I am capable of very little as a result of my experiences") and 11 (" It is hard for me to work or keep a job because of my experiences") to correlate would improve fit. As both items clearly relate to expectations around functioning, this residual correlation seemed appropriate. Modification indices also suggested that the residuals associated with item 11 and item 10 ("There is something about my personality that causes these experiences") correlate. However, there was no clear conceptual reason for this modification and this residual correlation was therefore left fixed to zero. Allowing the residuals for items seven and 11 to correlate led to a better fitting model (Model 4), scaled $\Delta X^2 = 24.51$, $p < .01$. Eight out of 60 standardized implied residuals for this model fell above .10. The parameters associated with this final model are reported in Table 5. All standardized loadings were above .40.

Model 4 produced reasonable but not excellent fit. Within this model the lowest factor loading was for item 11 (.46). As this item was also involved in the suggested residual correlations, it was felt that a simpler model may be obtainable by excluding item 11 (model

5). This final model demonstrated excellent fit with the data. Six out of 60 standardized implied residuals for this model fell above .10. The parameters associated with this model are reported in Table 5. All standardized loadings were above .40.

The largest Cook's distance for Models 4 and 5 was 0.0009, falling below the recommend cut-off of $4/N = 0.014$ for potential influential cases (Bollen & Jackman, 1990), suggesting that no individual cases were having an undue influence upon the models.

As the data for the present study came from two different sources (the EDIE2 and ACTION trial) there is a possibility that the factor structure being tested does not apply equally across both subsamples. Due to the small sample size, particularly in the ACTION sample (n = 63), it was not appropriate to conduct a multi-group CFA, or to test differential item functioning through other means, such as with a Multiple Indicators Multiple Causes (MIMIC) model (Finch, 2005). We instead conducted a sensitivity analysis by re-calculating the model with the ACTION subsample excluded. With the ACTION data excluded, model 5 remained the best fitting of those compared, although the overall fit of this model was marginally poorer, X^2 (51) = 96.24, CFI = .95, TLI = .93, RMSEA = .06 (.04 - .08). Factor loadings were similar, with all standardized loadings remaining above .40.

TABLE 5 AROUND HERE

Subscale Reliability and Validity

Three subscale scores were created by summing the items that loaded onto the three factors in Model 5.Internal consistencies and descriptive statistics for the three subscales are reported in Table 3. Internal reliability was acceptable for the ISD and NEA subscales but again low for the ES subscale (possibly due to the small number of items). The three subscales demonstrated moderate correlations with depressive symptoms (Table 6). Within a multiple regression analysis, each of the three subscales made significant independent contributions to the prediction of depressive symptoms (see semi-partial correlations in Table

3), accounting for 61% of the variance in symptoms, F (3, 285) = 54.43, p < .01. Residuals were homoscedastic and normally distributed. Equivalent findings emerged after one case with a large standardized residual (> 2.5) was dropped.

TABLE 6 AROUND HERE

Discussion

The three-factor model was the best fitting for the PBEQ, and demonstrated good overall fit. All three subscales made independent contributions in explaining depression scores. The current study leaves some uncertainty about how to treat the NEA subscale. Removal of item 11 led to an overall better fitting model, but there was no clear theoretical basis for how this item behaved within the model. Moreover, the internal reliability of the subscale is improved by keeping item 11. Future researchers should therefore be mindful of the behaviour of this item when using the PBEQ. A summary of the measure items and subscales is provided in Table 7.

TABLE 7 ABOUT HERE

A subset of the data from Study 2 were the same as the data used for the initial PCA in the study by Pyle et al (in press). CFA is quite a different analysis to PCA, varying in orientation (exploratory vs. confirmatory) and treatment of item variance (focus on total vs. common variance). Moreover, CFA has the particular strength of allowing multiple alternative models to be estimated and compared, and is typically seen as a subsequent step to exploratory approaches like PCA (Costello & Osborne, 2005). In this way the current study expands on the findings of Pyle et al (in press), and Pyle et al was an important study in developing the PBEQ and suggesting an initial factor structure. However, the re-use of these data are a limitation in that it raises the question of whether the identified results would generalize to other samples. The concept for the current study emerged after the acceptance

of the paper by Pyle et al (in press) and as such it was not possible to integrate these studies into a single paper.

The current psychosis sample consisted both of ARMS individuals and those with more established psychosis. Combining these subsamples had the advantage of allowing a test of the factor structure in a sample with a diverse range of psychotic experience. The 3-factor model was supported in the combined sample and in the ARMS sample on its own, but it was not possible to compare the fit of this model between these two groups. This means that some aspects of the model (e.g., item loadings) may differ between these groups and this could have implications for how the PBEQ is used (e.g., some items not used in certain groups). Further research is needed to confirm the factor structure in larger samples of individuals diagnosed with psychotic disorders and to compare factor models between groups.

General Discussion

The aim of the study was to test three theoretically informed models of the factor structure of the PBIllQ and PBEQ. A similar, novel three-factor structure was supported for both measures, including factors related to beliefs about internal shame/defectiveness, external shame and general negative appraisals of experiences. Shame is a complex emotional and cognitive construct (Gilbert, 2006). The current factors draw on the cognitive elements of shame, such as perceptions of inferiority relative to others, and build on previous theory concerning the distinction between internal and external forms of shame. The factors are consistent with the presence of shame/defectiveness type schema observed by other studies in individuals with psychosis and bipolar disorder (Addington & Tran, 2009; Fowke, Ross, & Ashcroft, 2012; Fowler et al., 2006). Importantly, however, since the two measures have been completed in different samples it is not known whether the same 3-factor structure would be replicated by both measures in the same sample.

It should be noted that the ES subscale had low internal reliability for both measures, likely related to the small number of items featured in this subscale. As such, clinicians and investigators should be wary in using this specific subscale, and mindful of its psychometric properties. Adding additional items to the scale that relate to this specific factor would be one way to enhance the properties of this subscale, and could be pursued in future research.

Several studies have supported the notion that individuals' beliefs about their psychotic experiences are an important determinate of subsequent distress and affective disturbance, including depression (Acosta et al., 2013; Birchwood et al., 1993; Karatzias et al., 2007; Stainsby et al., 2010; White et al., 2007). The current study is consistent with this notion, with the ISD and NAE factors demonstrating moderate independent relationships with depressive symptoms. As the ISD captures cognitive elements of shame, the correlation with depression is also consistent with a meta-analysis supporting close links between shame and depression (Kim, Thibodeau, & Jorgensen, 2011).

The Illness Perceptions Questionnaire (IPQ) has also been modified several times for use in those with psychosis (Lobban, Barrowclough, & Jones, 2004; Marcus et al., 2014; Watson et al., 2006). Although the modified IPQ shares some similar content with the PBIllQ and PBEQ, both considering causes and control over experiences, there are also some notable differences. In particular the modified IPQ has few or no items relating to a sense of dysfunction, social approval, shame or stigma around symptoms. As such the modified IPQ and the PBEQ/PBIllQ likely capture overlapping but also distinct belief sets. Future research should consider contrasting these measures and examining whether they could be effectively combined to better assess negative beliefs.

Within clinical practice the PBEQ and PBIllQ may provide helpful tools for identifying individuals' beliefs about their psychosis or bipolar disorder. The PBEQ may be suitable for a broader population than the PBIllQ, from which it is based, due to the changes

in item wording. The PBEQ also demonstrated better fit in the current study, although this may also reflect differences in the sample. Information obtained with these measures could be used to inform formulations of clients' difficulties, either for use within individual therapy or with staff teams to inform service planning. The current study is important in providing evidence about the putative belief sets that the PBEQ/PBIlQ items map onto. This information is important in scoring these measures but also in their interpretation. Hence, when formulating a client's difficulties it may be helpful to distinguish between general negative appraisals about their mental health and those beliefs associated with perceived dysfunction or internal shame and consider these as separate components within the formulation. Different profiles on the PBEQ or PBIllQ could inform different interventions. An individual who reports general negative appraisals and expectations about their psychosis or bipolar disorder may benefit from structured behavioural interventions designed to improve their expectations about what they can achieve. In contrast, in an individual where a sense of internal shame and defectiveness is substantial, work on challenging these self-perceptions and enhancing positive affect may first be necessary. As noted, the PBIllQ/PBEQ have been used in the context of clinical trials where they have the potential to provide insight into a possible mechanism of change for cognitively-orientated therapies.

The current findings also raise the question of whether the results of previous studies and trials using the PBEQ or PBIllQ, which are based upon subscales other than the three supported by the current study, should be revised. It is possible that the use of the 3-factor structure supported in this study would lead to different results. Inasmuch as the 3-factor model appears to generalise to these other samples researchers may wish to re-evaluate their findings. Notably, two recent studies using the PBEQ (Pyle et al., in press; Stowkowy et al., in press) take different approaches to deriving scores (2 versus 4 subscales) limiting

comparability. It is hoped the current study will lead to a greater consistency in how this scale is used in future.

References

Acosta, F. J., Aguilar, E. J., Cejas, M. R., & Gracia, R. (2013). Beliefs about illness and their relationship with hopelessness, depression, insight and suicide attempt in schizophrenia. *Psychiatria Danubina, 25*, 49-54. doi: http://www.hdbp.org/psychiatria_danubina/pdf/dnb_vol25_no1/dnb_vol25_no1_49.pdf

Addington, J., & Tran, L. (2009). Using the brief core schema scales with individuals at clinical high risk of psychosis. *Behavioural and Cognitive Psychotherapy, 37*, 227-231. doi: 10.1017/S1352465809005116

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text Revision ed.). Washington, DC: Author.

Beck, A. T., Guth, D., Steer, R. A., & Ball, R. (1997). Screening for major depression disorders in medical inpatients with the Beck Depression Inventory for Primary Care. *Behaviour Research & Therapy, 35*, 785-791. doi: 10.1016/S0005-7967(97)00025-9

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.

Birchwood, M., Jackson, C., Brunet, K., Holden, J., & Barton, K. (2012). Personal beliefs about illness questionnaire-revised (PBIQ-R): reliability and validation in a first episode sample. *British Journal of Clinical Psychology, 51*, 448-458. doi: 10.1111/j.2044-8260.2012.02040.x

Birchwood, M., Mason, R., MacMillan, F., & Healy, J. (1993). Depression, demoralization and control over psychotic illness: A comparison of depressed and non-depressed patients with a chronic psychosis. *Psycholoigcal Medicine, 23*, 387-395.

Bollen, K. A., & Jackman, R. W. (1990). Regression diagnostics: An expository treatment of outliers and influential cases. In J. Fox & J. S. Long (Eds.), *Modern methods of data analysis* Newbury Park, CA: Sage.

Byrne, B. (2001). *Structural equation modelling with AMOS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.

Byrne, R., & Morrison, A. P. (2010). Young people at risk of psychosis: a user-led exploration of interpersonal relationships and communication of psychological difficulties. *Early Intervention in Psychiatry, 4*, 162-168. doi: 10.1111/j.1751-7893.2010.00171.x

Chalmers, P. (2013). faoutlier: Influential case detection methods for factor analysis and SEM. R package version 0.4. https://github.com/philchalmers/faoutlierdoi: https://github.com/philchalmers/faoutlier

Clatworthy, J., Bowskill, R., Rank, T., Parham, R., & Horne, R. (2007). Adherence to medication in bipolar disorder: A qualitative study exploring the role of patients' beliefs about the condition and its treatment. *Bipolar Disorders, 9*(6), 656-664. doi: 10.1111/j.1399-5618.2007.00434.x

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*, 1-9. doi: http://pareonline.net/getvn.asp?v=10&n=7

Elnazer, H. Y., Brannac, A., & Kingdon, D. (2014). *Beliefs about bipolar affective disorder assessment and clinical implications*. Paper presented at the Royal College of Psychiatry International COngress, London.

Finch, H. (2005). The MIMIC Model as a method for detecting DIF: A comparison with Mantel-Haenszel, SIBTEST, and the IRT likelihood ratio. *Applied Psychological Measurment, 29*, 278-295. doi: 10.1177/0146621605275728

First, M., Spitzer, R., Gibbon, M., & Williams, B. (1996). *Structured Clinical Interview for DSM-IV Axis I Disorders, Clinician Version (SCID-CV)*. Washington, D.C.: American Psychiatric Press, Inc.

Fowke, A., Ross, S., & Ashcroft, K. (2012). Childhood maltreatment and internalized shame in adults with a diagnosis of bipolar disorder. *Clinical Psychology & Psychotherapy, 19*, 450-457. doi: 10.1002/cpp.752

Fowler, D., Freeman, D., Smith, B., Kuipers, E., Bebbington, P., Bashforth, H., . . . Garety, P. (2006). The Brief Core Schema Scales (BCSS): Psychometric properties and associations with paranoia and grandiosity in non-clinical and psychosis samples. *Psychological Medicine, 36*, 749-759. doi: 10.1017/S0033291706007355

Gilbert, P. (2006). A biopsychosocial and evolutionary approach to formulation with a special focus on shame. In N. Tarrier (Ed.), *Case formulation in cognitive behaviour therapy: The treatment of challenging and complex cases*. London: Routledge.

Green, G., Hayes, C., Dickinson, D., Whittaker, A., & Gilheany, B. (2003). A mental health service users perspective to stigmatisation. *Journal of Mental Health, 12*, 223-234. doi: 10.1080/0963823031000118212

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1-55. doi: 10.1080/10705519909540118

Karatzias, T., Gumley, A., Power, K., & O'Grady, M. (2007). Illness appraisals and self-esteem as correlates of anxiety and affective comorbid disorders in schizophrenia. *Comprehensive Psychiatry, 48*, 371-375. doi: 10.1016/j.comppsych.2007.02.005

Kay, S., Fiszbein, A., & Opler, L. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin, 13*, 261 - 276. doi: 10.1093/schbul/13.2.261

Kim, S., Thibodeau, R., & Jorgensen, R. S. (2011). Shame, guilt, and depressive symptoms: A meta-analytic review. *Psycholoigcal Bulletin, 137*, 68-96. doi: 10.1037/a0021466

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: The Guilford Press.

Kumar, G., Rissmiller, D. J., Steer, R. A., & Beck, A. T. (2006). Mean Beck depression Inventory-II total scores by type of bipolar episode. *Psychological Reports, 98*, 836-840. doi: 10.2466/pr0.98.3.836-840

Lobban, F., Barrowclough, C., & Jones, S. (2004). The impact of beliefs about mental health problems and coping on outcome in schizophrenia. *Psychological Medicine, 34*(7), 1165-1176. doi: 10.1017/s003329170400203x

Marcus, E., Garety, P., Weinman, J., Emsley, R., Dunn, G., Bebbington, P., . . . Jolley, S. (2014). A pilot validation of a modified Illness Perceptions Questionnaire designed to predict response to cognitive therapy for psychosis. *Journal of Behavior Therapy and Experimental Psychiatry, 45*(4), 459-466. doi: http://dx.doi.org/10.1016/j.jbtep.2014.06.003

Morrison, A. P., French, P., Stewart, S. L. K., Birchwood, M., Fowler, D., Gumley, A. I., . . . Dunn, G. (2012). Early detection and intervention evaluation for people at risk of psychosis: Multisite randomised controlled trial. *British Medical Journal, 344*, 2233-2233. doi: 10.1136/bmj.e2233

Morrison, A. P., Stewart, S. L. K., French, P., Bentall, R. P., Birchwood, M., Byrne, R., . . . Dunn, G. (2011). Early detection and intervention evaluation for people at high-risk of psychosis-2 (EDIE-2): Trial rationale, design and baseline characteristics. *Early Intervention in Psychiatry, 5*, 24-32. doi: 10.1111/j.1751-7893.2010.00254.x

Morrison, A. P., Wardle, M., Hutton, P., Davies, L., Dunn, G., Brabban, A., . . . Turkington, D. (2013). Assessing cognitive therapy instead of neuroleptics: Rationale, study design and sample characteristics of the ACTION trial. *Psychosis: Psychological, Social and Integrative Approaches, 5*, 82-92. doi: 10.1080/17522439.2012.756539

O'Leary-Kelly, S. W., & Vokurka, R. J. (1998). The empirical assessment of construct validity. *Journal of Operations Management, 16*, 387-405. doi: 10.1016/S0272-6963(98)00020-5

Pyle, M., Stewart, S. L. K., French, P., Byrne, R., Patterson, P., Gumley, A., . . . Morrison, A. P. (in press). Internalized stigma, emotional dysfunction and unusual experiences in young people at risk of psychosis. *Early Intervention in Psychiatry*. doi: 10.1111/eip.12098

R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistcial Software, 48*, 1-36. doi: http://www.jstatsoft.org/v48/i02/

Sayre, J. (2000). The patient's diagnosis: Explanatory models of mental illness. *Qualitative Health Research, 10*, 71-83. doi: 10.1177/104973200129118255

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*, 147-177. doi: 10.1037//1082-989X.7.2.147

Scott, J., & Pope, M. (2002). Nonadherence with mood stabilizers: prevalence and predictors. *Journal of Clinical Psychiatry, 63*(5), 384-390. doi: 10.4088/JCP.v63n0502

Stainsby, M., Sapochnik, M., Bledin, K., & Mason, O. J. (2010). Are attitudes and beliefs about symptoms more important than symptom severity in recovery from psychosis? *Psychosis, 2*, 41-49. doi: 10.1080/17522430903144386

Steer, R. A., Cavalieri, T. A., Leonard, D. M., & Beck, A. T. (1999). Use of the Beck depression inventory for primary care to screen for major depression disorders. *General Hospital Psychiatry, 21*, 106-111. doi: http://dx.doi.org/10.1016/S0163-8343(98)00070-X

Stowkowy, J., Perkins, D. O., Woods, S. W., Nyman, K., & Addington, J. (2014). Personal Beliefs about Experiences in those at Clinical High Risk for Psychosis. *Behav Cogn Psychother, FirstView*, 1-7. doi: doi:10.1017/S1352465814000307

Stowkowy, J., Perkins, D. O., Woods, S. W., Nyman, K., & Addington, J. (in press). Personal beliefs about experiences in those at clinical high risk for psychosis. *Behavioural Cognitive Psychotherapy*, 1-7. doi: 10.1017/s1352465814000307

Taylor, H. E., Stewart, S. L. K., Dunn, G., Parker, S., Bentall, R. P., Birchwood, M., & Morrison, A. P. (in press). Psychopathology and affect dysregulation across the continuum of psychosis: a multiple comparison group study. *Early Intervention in Psychiatry*. doi: 10.1111/eip.12064

Wang, Y. P., & Gorenstein, C. (2013). Psychometric properties of the Beck Depression Inventory-II: A comprehensive review. *Revista Brasileira de Psiquiatria, 35*, 416-431. doi: 10.1590/1516-4446-2012-1048

Watson, P. W., Garety, P. A., Weinman, J., Dunn, G., Bebbington, P. E., Fowler, D., . . . Kuipers, E. (2006). Emotional dysfunction in schizophrenia spectrum psychosis: The role of illness perceptions. *Psychological Medicine, 36*, 761-770. doi: 10.1017/S0033291706007458

White, R. G., McCleery, M., Gumley, A. I., & Mulholland, C. (2007). Hopelessness in schizophrenia: The impact of symptoms and beliefs about illness. *Journal of Nervous and Mental Disease, 195*, 968-975. doi: 10.1097/NMD.0b013e31815c1a1d

Yung, A. R., Yuen, H. P., McGorry, P. D., Phillips, L. J., Kelly, D., Dell'Olio, M., . . . Buckby, J. (2005). Mapping the onset of psychosis: The Comprehensive Assessment of At-Risk Mental States. *Austalian an New Zealand Journal of Psychiatry, 39*, 964–971. doi: 10.1111/j.1440-1614.2005.01714.x

Table 1: Fit Indices from CFA for the PBIllQ

Model	Model fit				
	Scaled X^2	*df*	CFI	TLI	RMSEA
Model 1: 2 factor	412.52	103	.80	.77	.13 (.11-.14)
Model 2: 3 factor	271.20	101	.89	.87	.09 (.08-.11)
Model 3: 5 factor	262.45	94	.89	.86	.10 (.08-.11)
Model 4: 3 factor with 1 correlated error	209.56	100	.93	.92	.08 (.06-.09)
Model 5: 3 factor with 2 correlated errors	193.70	99	.94	.93	.07 (.06-.09)

Note: All statistics based on robust ordinal estimation; Model 3 did not converge normally, producing an inadmissible factor correlation; CFI = Confirmatory Fit Index; RMSEA = Root Mean Squared Residual; TLI = Tucker Lewis Fit Index

Table 2: Standardized and Unstandardized Parameter in Final Model for the PBIIlQ

Parameter	Model 5 Unstandardized estimate	Standardized estimate
Factor loadings		
Internal shame/defectiveness (ISD)		
PBIIlQ 13	1.00	.46
PBIIlQ 12	1.02	.47
PBIIlQ 14	0.88	.40
PBIIlQ 16	1.38	.63
PBIIlQ 6	1.61	.74
PBIIlQ 10	1.78	.81
External shame (ES)		
PBIIlQ 15	1.00	.54
PBIIlQ 3	1.46	.79
Negative expectations/appraisals (NEA)		
PBIIlQ 7	1.00	.79
PBIIlQ 4	0.90	.71
PBIIlQ 8	0.75	.59
PBIIlQ 2	0.92	.73
PBIIlQ11	0.88	.69
PBIIlQ 1	0.54	.43
PBIIlQ 5	0.64	.51
PBIIlQ 9	0.83	.66
Factor covariances & correlations		
ISD – ES	0.13	.54
ISD - NEA	0.23	.63
ES - NEA	0.33	.76

All parameters significant, $p < .05$

Table 3

Descriptive Statistics, Internal Reliability and Correlations for three PBIllQ Subscales

	Descriptives			Zero-order Correlations	Multiple Regression	
Variable	Alpha	Mean	SD	Depression	*B*	Semi-partial correlations
ISD	.70	12.82	3.34	.52*	1.19*	.26
ES	.53	5.27	1.47	.36*	0.25	.02
NEA	.81	19.02	4.48	.62*	1.46*	.40

* $p < .01$; *n* varies between 190 – 201 due to missing data

Table 4: Fit Indices from CFA for the PBEQ

	Model fit				
Model	Scaled X^2	*df*	CFI	TLI	RMSEA
Model 1: 2 factor	325.23	64	.81	.77	.12 (.11-.13)
Model 2: 3 factor	169.99	62	.92	.90	.08 (.06-.09)
Model 3: 4 factor	210.61	59	.89	.85	.10 (.08-.11)
Model 4: 3 factor with correlated errors	148.51	61	.94	.92	.07 (.06-.09)
Model 5: 3 factor with item 11 excluded	98.43	51	.96	.95	.06 (.04-.07)

Note: All statistics based on robust ordinal estimation; CFI = Confirmatory Fit Index; RMSEA = Root Mean Squared Residual; TLI = Tucker Lewis Fit Index

Table 5: Standardized and Unstandardized Parameters for Models 4 and 5 with the PBEQ

	Model 4		Model 5	
Parameter	Unstandardized estimate	Standardized estimate	Unstandardized estimate	Standardized estimate
Factor loadings				
Internal shame/defectiveness (ISD)				
PBEQ 4	1.00	.56	1.00	.57
PBEQ 6	0.76	.43	0.77	.44
PBEQ 10	1.13	.64	1.05	.60
PBEQ 13	1.22	.69	1.23	.70
PBEQ 2	1.23	.69	1.24	.70
PBEQ 8	1.04	.58	1.03	.59
External shame (ES)				
PBEQ 12	1.00	.63	1.00	.64
PBEQ 3	1.33	.83	1.29	.82
Negative expectations/appraisals (NEA)				
PBEQ 5	1.00	.81	1.00	.80
PBEQ 7	0.82	.66	0.83	.66
PBEQ 1	0.65	.53	0.68	.54
PBEQ 9	0.75	.61	0.78	.62
PBEQ11	0.57	.46		-
Factor covariances & correlations				
ISD – ES	0.15	.42	0.15	.42
ISD - NEA	0.30	.65	0.29	.64
ES - NEA	0.21	.42	0.22	.44

All parameters significant, $p < .05$

Table 6

Descriptive Statistics, Internal Reliability and Correlations for three PBEQ Subscales

	Descriptives			Zero-order Correlations	Multiple Regression	
Variable	Alpha	Mean	SD	Depression	*B*[a]	Semi-partial correlations[a]
ISD	.72	15.66	3.19	.49*	.41*	.25
ES	.61	5.73	1.37	.35*	.62*	.18
NEA	.68	10.77	2.33	.50*	.59*	.27
NEA (with item 11 included)	.71	13.48	2.85	.50*	.49*	.28

* $p < .01$; [a] results for the ISD and ES derived from model were the NEA variable includes item 11. *n* varies between 286 – 303 due to missing data

Table 7

Summary of PBEQ and PBIllQ items and subscales

Item	Subscale
PBIllQ	
1. I will always need to be cared for by professional medical staff	NEA
2. My illness frightens me	NEA
3. I am embarrassed by my illness	ES
4. I am capable of very little as a result of my illness	NEA
5. Because of my illness I have to rely on psychiatric services	NEA
6. There must always have been something wrong with me to have caused my illness	ISD
7. I find it difficult to cope with my current symptoms	NEA
8. My illness is too delicate/brittle for me to work or keep a job	NEA
9. I know when I'm relapsing but I can't do anything about it	NEA
10. My illness is a judgement on me	ISD
11. I am powerless to influence or control my illness	NEA
12. I am fundamentally normal, my illness is like any other	ISD
13. Society needs to keep people with my illness apart from everyone else	ISD
14. There must be something about my personality that causes me to be what I am	ISD
15. I can talk to most people about my illness	ES
16. There is something strange about me that causes my illness	ISD
PBEQ	
1. My experiences frighten me…	NEA
2. There must always have been something wrong with me as a person (to have caused these experiences)…	ISD
3. I am embarrassed to talk about my experiences…	ES

4.	My experiences may mean that I should be kept away from others…	ISD
5.	I find it difficult to cope with my current experiences…	NEA
6.	I am fundamentally normal, my experiences are like any other…	ISD
7.	I am capable of very little as a result of my experiences…	NEA
8.	My experiences are a judgement on me…	ISD
9.	I am powerless to influence or control my experiences…	NEA
10.	There is something about my personality that causes these experiences…	ISD
11.	It is hard for me to work or keep a job because of my experiences…	NEA
12.	I can talk to most people about my experiences…	ES
13.	There is something strange about me which is responsible for these experiences…	ISD

ES = External Shame subscale; ISD = Internal Shame/Defectiveness subscale; NEA = Negative Appraisal of Experiences subscale; PBEQ = Personal Beliefs about Experiences Questionnaire; PBIllQ = Personal Beliefs about Illness Questionnaire